NORTH-HOLLAND

On a Class of Block Toeplitz Matrices

Robert L. Ellis
Department of Mathematics
University of Maryland
College Park, Maryland 20740

Israel Gohberg
Raymond and Beverly Sackler
Faculty of Exact Sciences
Tel Aviv University
Ramat Aviv 69978, Israel

and

David C. Lay
Department of Mathematics
University of Maryland
College Park, Maryland 20740

Submitted by Harm Bart

ABSTRACT

In this paper we consider a class of matrices, each of which is the sum of an identity matrix and a self-adjoint block Toeplitz matrix that has a symmetric band of zero blocks. A number of general results for block Toeplitz matrices are specialized to this class.

1. INTRODUCTION

In this paper we study block Toeplitz matrices of the special form

$$\begin{pmatrix} I & G \\ G^* & I \end{pmatrix}, \tag{1.1}$$

LINEAR ALGEBRA AND ITS APPLICATIONS 241–243:225–245 (1996)

655 Avenue of the Americas, New York, NY 10010

0024-3795/96/$15.00
SSDI 0024-3795(95)00260-X

where I is an identity matrix and G is a block upper-triangular Toeplitz matrix. Thus

$$G = \begin{pmatrix} g_n & \cdots & g_0 \\ & \ddots & \vdots \\ 0 & & g_n \end{pmatrix},$$

where each g_k $(0 \leqslant k \leqslant n)$ is a $p \times p$ complex matrix. Matrices in the form (1.1) arise in the problem of orthogonalizing certain matrix-valued functions on the unit circle [3]. The results here are used in our work in [3].

In the extensive literature that already exists on block Toeplitz matrices, a special role is played by the equation

$$T\begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} I \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{1.2}$$

Here T is an $(m+1) \times (m+1)$ block Toeplitz matrix with $p \times p$ blocks, I is the $p \times p$ identity matrix, and $x_0, \dots, x_m$ are $p \times p$ matrices. Associated with $x_0, \dots, x_m$ is the polynomial f defined by

$$f(z) = x_0 z^m + x_1 z^{m-1} + \cdots + x_m. \tag{1.3}$$

When T has the form in (1.1), it is natural to rewrite Eq. (1.2) as

$$\begin{pmatrix} I & G \\ G^* & I \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} e_1 \\ 0 \end{pmatrix}, \tag{1.4}$$

where $e_1 = (I, 0, \dots, 0)^T$ and a and b have the form

$$a = (\alpha_0, \dots, \alpha_n)^T \qquad \text{and} \qquad b = (\beta_{-n}, \dots, \beta_0)^T.$$

Here the superscript T indicates the transpose of a matrix. In this case the corresponding functions

$$\alpha(z) = \alpha_0 + \alpha_1 z + \cdots + \alpha_n z^n \tag{1.5}$$

and

$$\beta(z) = \beta_0 + \beta_{-1}z^{-1} + \cdots + \beta_{-n}z^{-n} \tag{1.6}$$

are of interest in addition to the polynomial in (1.3).

Our goal is to investigate the relationships among G, the vectors a and b, the functions α and β, and the polynomial f. Of particular interest is the distribution of zeros of the functions α and β. The scalar case ($p = 1$) has already been explored in [2, Sect. 6]. However, in the matrix case to be studied here, additional difficulties arise, especially in the investigation of the zeros of α and β.

The first result in Section 2 shows that if G, a, and b satisfy (1.4), then α and β satisfy the identity

$$\alpha(z)^* \alpha(z) - \beta(z)^* \beta(z) = \alpha_0 \qquad (|z| = 1) \tag{1.7}$$

and α_0 is self-adjoint. Furthermore, if α and β are matrix-valued functions of the form (1.5) and (1.6), with α_0 self-adjoint and invertible, and if (1.7) holds, then there exists a unique block upper-triangular Toeplitz matrix G such that (1.4) holds. In fact,

$$G = -\begin{pmatrix} \alpha_0^* & \cdots & \alpha_n^* \\ & \ddots & \vdots \\ 0 & & \alpha_0^* \end{pmatrix}^{-1} \begin{pmatrix} \beta_{-n}^* & \cdots & \beta_0^* \\ & \ddots & \vdots \\ 0 & & \beta_{-n}^* \end{pmatrix}.$$

A related result gives necessary and sufficient conditions for the existence of a matrix G as above to satisfy both (1.4) and the related equation

$$\begin{pmatrix} I & G \\ G^* & I \end{pmatrix}\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ \hat{e}_1 \end{pmatrix}$$

for specified c and d, where $\hat{e}_1 = (0, \ldots, 0, I)^T$.

Theorem 2.3 concerns the zeros of the determinants of the matrix-valued functions α, β and f, assuming that a, b, and G satisfy (1.4). If α_0 is positive definite, then $\alpha(z)$ and $f(z)$ are invertible for $|z| = 1$, and the number of zeros of $\det \alpha$ inside the unit circle equals both the number of zeros of $\det f$ outside the unit circle and the number of negative eigenvalues of the matrix

$$T_1 = \begin{pmatrix} I & G_1 \\ G_1^* & I \end{pmatrix}. \tag{1.8}$$

where G_1 is the $n \times n$ matrix given by

$$G_1 = \begin{pmatrix} g_n & \cdots & g_1 \\ & \ddots & \vdots \\ 0 & & g_n \end{pmatrix}.$$

If α_0 is negative definite, then $\beta(z)$ and $f(z)$ are invertible for $|z| = 1$, and the number of zeros of $\det \beta$ outside the unit circle equals both the number of zeros of $\det f$ inside the unit circle and the number of negative eigenvalues of the matrix T_1 in (1.8).

2. THEOREMS AND PROOFS

In this section we state and prove the theorems that were described in the Introduction.

THEOREM 2.1. *Let $a = (\alpha_0, \dots, \alpha_n)^T$ and $b = (\beta_{-n}, \dots, \beta_0)^T$ be given vectors with $p \times p$ matrix entries and with α_0 invertible. Then there exists a block upper-triangular Toeplitz matrix G such that*

$$\begin{pmatrix} I & G \\ G^* & I \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} e_1 \\ 0 \end{pmatrix} \tag{2.1}$$

if and only if α_0 is self-adjoint and the following equation holds:

$$\begin{pmatrix} \alpha_0^* & \cdots & \alpha_n^* \\ & \ddots & \vdots \\ 0 & & \alpha_0^* \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix} - \begin{pmatrix} \beta_{-n}^* & \cdots & \beta_0^* \\ & \ddots & \vdots \\ 0 & & \beta_{-n}^* \end{pmatrix} \begin{pmatrix} \beta_{-n} \\ \vdots \\ \beta_0 \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{2.2}$$

The matrix G is unique and is given by

$$G = - \begin{pmatrix} \alpha_0^* & \cdots & \alpha_n^* \\ & \ddots & \vdots \\ 0 & & \alpha_0^* \end{pmatrix}^{-1} \begin{pmatrix} \beta_{-n}^* & \cdots & \beta_0^* \\ & \ddots & \vdots \\ 0 & & \beta_{-n}^* \end{pmatrix}.$$

Equation (2.2) *can be rewritten in the form*

$$\alpha(z)^* \alpha(z) - \beta(z)^* \beta(z) = \alpha_0 \qquad (|z| = 1), \tag{2.3}$$

where

$$\alpha(z) = \alpha_0 + \alpha_1 z + \cdots + \alpha_n z^n$$

and

$$\beta(z) = \beta_0 + \beta_{-1} z^{-1} + \cdots + \beta_{-n} z^{-n}.$$

Proof. Suppose a block upper-triangular Toeplitz matrix

$$G = \begin{pmatrix} g_n & \cdots & g_0 \\ & \ddots & \vdots \\ 0 & & g_n \end{pmatrix}$$

exists so that (2.1) is satisfied. Then

$$a + Gb = e_1 \tag{2.4}$$

and

$$G^* a + b = 0. \tag{2.5}$$

We may rewrite (2.5) as

$$a^* G + b^* = 0. \tag{2.6}$$

In matrix form (2.4) and (2.6) become

$$\begin{pmatrix} g_n & \cdots & g_0 \\ & \ddots & \vdots \\ 0 & & g_n \end{pmatrix} \begin{pmatrix} \beta_{-n} \\ \vdots \\ \beta_0 \end{pmatrix} = \begin{pmatrix} I - \alpha_0 \\ -\alpha_1 \\ \vdots \\ -\alpha_n \end{pmatrix} \tag{2.7}$$

and

$$\begin{pmatrix} \alpha_0^* & \cdots & \alpha_n^* \end{pmatrix} \begin{pmatrix} g_n & \cdots & g_0 \\ & \ddots & \vdots \\ 0 & & g_n \end{pmatrix} = \begin{pmatrix} -\beta_{-n}^* & \cdots & -\beta_0^* \end{pmatrix}. \quad (2.8)$$

Equations (2.7) and (2.8) are equivalent to

$$\begin{pmatrix} g_n & \cdots & g_0 \\ & \ddots & \vdots \\ 0 & & g_n \end{pmatrix} \begin{pmatrix} \beta_0 & \cdots & \beta_{-n} \\ & \ddots & \vdots \\ 0 & & \beta_0 \end{pmatrix} = \begin{pmatrix} -\alpha_n & \cdots & -\alpha_1 & I - \alpha_0 \\ & \ddots & & -\alpha_1 \\ & & & \vdots \\ 0 & & & -\alpha_n \end{pmatrix} \quad (2.9)$$

and

$$\begin{pmatrix} \alpha_0^* & \cdots & \alpha_n^* \\ & \ddots & \vdots \\ 0 & & \alpha_0^* \end{pmatrix} \begin{pmatrix} g_n & \cdots & g_0 \\ & \ddots & \vdots \\ 0 & & g_n \end{pmatrix} = \begin{pmatrix} -\beta_{-n}^* & \cdots & -\beta_0^* \\ & \ddots & \vdots \\ 0 & & -\beta_{-n}^* \end{pmatrix}. \quad (2.10)$$

Multiplying (2.9) on the left by

$$\begin{pmatrix} \alpha_0^* & \cdots & \alpha_n^* \\ & \ddots & \vdots \\ 0 & & \alpha_0^* \end{pmatrix}$$

and (2.10) on the right by

$$\begin{pmatrix} \beta_0 & \cdots & \beta_{-n} \\ & \ddots & \vdots \\ 0 & & \beta_0 \end{pmatrix}$$

and subtracting the resulting equations, we obtain

$$\begin{pmatrix} \alpha_0^* & \alpha_1^* & \cdots & \alpha_n^* \\ & \alpha_0^* & \cdots & \alpha_{n-1}^* \\ & & \ddots & \vdots \\ 0 & & & \alpha_0^* \end{pmatrix} \begin{pmatrix} -\alpha_n & \cdots & -\alpha_1 & I-\alpha_0 \\ & \ddots & & -\alpha_1 \\ & & & \vdots \\ 0 & & & -\alpha_n \end{pmatrix}$$

$$-\begin{pmatrix} -\beta_{-n}^* & \cdots & -\beta_0^* \\ & \ddots & \vdots \\ 0 & & -\beta_{-n}^* \end{pmatrix} \begin{pmatrix} \beta_0 & \cdots & \beta_{-n} \\ & \ddots & \vdots \\ 0 & & \beta_0 \end{pmatrix} = 0.$$

Rearranging yields

$$\begin{pmatrix} \alpha_0^* & \cdots & \alpha_n^* \\ & \ddots & \vdots \\ 0 & & \alpha_0^* \end{pmatrix} \begin{pmatrix} \alpha_n & \cdots & \alpha_0 \\ & \ddots & \vdots \\ 0 & & \alpha_n \end{pmatrix}$$

$$-\begin{pmatrix} \beta_{-n}^* & \cdots & \beta_0^* \\ & \ddots & \vdots \\ 0 & & \beta_{-n}^* \end{pmatrix} \begin{pmatrix} \beta_0 & \cdots & \beta_{-n} \\ & \ddots & \vdots \\ 0 & & \beta_0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & \cdots & \alpha_0^* \\ 0 & & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

Equating the last columns of both sides of this equation, we obtain

$$\begin{pmatrix} \alpha_0^* & \cdots & \alpha_n^* \\ & \ddots & \vdots \\ 0 & & \alpha_0^* \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix} - \begin{pmatrix} \beta_{-n}^* & \cdots & \beta_0^* \\ & \ddots & \vdots \\ 0 & & \beta_{-n}^* \end{pmatrix} \begin{pmatrix} \beta_{-n} \\ \vdots \\ \beta_0 \end{pmatrix} = \begin{pmatrix} \alpha_0^* \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \quad (2.11)$$

Now observe that for $|z| = 1$,

$$\begin{aligned} &\alpha(z)^* \alpha(z) - \beta(z)^* \beta(z) \\ &\quad = \left(\alpha_0^* + \alpha_1^* z^{-1} + \cdots + \alpha_n^* z^{-n}\right)\left(\alpha_0 + \alpha_1 z + \cdots + \alpha_n z^n\right) \\ &\qquad - \left(\beta_0^* + \beta_{-1}^* z + \cdots + \beta_{-n}^* z^n\right)\left(\beta_0 + \beta_{-1} z^{-1} + \cdots + \beta_{-n} z^{-n}\right). \end{aligned} \tag{2.12}$$

Since $\alpha(z)^*\alpha(z) - \beta(z)^*\beta(z)$ is self-adjoint, it follows that the coefficient of z^{-k} in (2.12) is the adjoint of the coefficient of z^k for $-n \leqslant k \leqslant n$. It is easy to check that (2.11) is equivalent to the equality of the coefficients of the nonnegative powers of z in the functions $\alpha(z)^*\alpha(z) - \beta^*(z)\beta(z)$ and α_0^*. Thus (2.11) is equivalent to

$$\alpha(z)^* \alpha(z) - \beta(z)^* \beta(z) = \alpha_0^* \qquad (|z| = 1).$$

This implies that α_0 is self-adjoint, so that (2.2) and (2.3) hold.

Conversely, suppose that α_0 is self-adjoint and that (2.2) holds. Since α_0 is invertible, we may define

$$\begin{pmatrix} g_0 \\ \vdots \\ g_n \end{pmatrix} = -\begin{pmatrix} \alpha_0^* & \cdots & \alpha_n^* \\ & \ddots & \vdots \\ 0 & & \alpha_0^* \end{pmatrix}^{-1} \begin{pmatrix} \beta_0^* \\ \vdots \\ \beta_{-n}^* \end{pmatrix}$$

and let

$$G = \begin{pmatrix} g_n & \cdots & g_0 \\ & \ddots & \vdots \\ 0 & & g_n \end{pmatrix}.$$

Then

$$\begin{pmatrix} \alpha_0^* & \cdots & \alpha_n^* \\ & \ddots & \vdots \\ 0 & & \alpha_0^* \end{pmatrix} \begin{pmatrix} g_0 \\ \vdots \\ g_n \end{pmatrix} = -\begin{pmatrix} \beta_0^* \\ \vdots \\ \beta_{-n}^* \end{pmatrix},$$

which is equivalent to

$$\begin{pmatrix} \alpha_0^* & \cdots & \alpha_n^* \\ & \ddots & \vdots \\ 0 & & \alpha_0^* \end{pmatrix} \begin{pmatrix} g_n & \cdots & g_0 \\ & \ddots & \vdots \\ 0 & & g_n \end{pmatrix} = - \begin{pmatrix} \beta_{-n}^* & \cdots & \beta_0^* \\ & \ddots & \vdots \\ 0 & & \beta_{-n}^* \end{pmatrix}. \tag{2.13}$$

Taking the adjoint of this equation and equating the first columns of both sides of the resulting equation, we obtain

$$G^*a + b = 0. \tag{2.14}$$

Using first (2.13) and then (2.2), we have

$$\begin{aligned} a + Gb &= \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix} - \begin{pmatrix} \alpha_0^* & \cdots & \alpha_n^* \\ & \ddots & \vdots \\ 0 & & \alpha_0^* \end{pmatrix}^{-1} \begin{pmatrix} \beta_{-n}^* & \cdots & \beta_0^* \\ & \ddots & \vdots \\ 0 & & \beta_{-n}^* \end{pmatrix} \begin{pmatrix} \beta_{-n} \\ \vdots \\ \beta_0 \end{pmatrix} \\ &= \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix} - \begin{pmatrix} \alpha_0^* & \cdots & \alpha_n^* \\ & \ddots & \vdots \\ 0 & & \alpha_0^* \end{pmatrix}^{-1} \\ &\quad \times \left[\begin{pmatrix} \alpha_0^* & \cdots & \alpha_n^* \\ & \ddots & \vdots \\ 0 & & \alpha_0^* \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix} - \begin{pmatrix} \alpha_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right] \\ &= \begin{pmatrix} \alpha_0^* & \cdots & \alpha_n^* \\ & \ddots & \vdots \\ 0 & & \alpha_0^* \end{pmatrix}^{-1} \begin{pmatrix} \alpha_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \end{aligned}$$

so that

$$a + Gb = e_1. \tag{2.15}$$

Equations (2.14) and (2.15) imply that (2.1) is satisfied by G. The uniqueness of G is a consequence of (2.10). ■

The next goal is to prove a theorem related to a theorem of Krein [6] on the zeros of orthogonal polynomials and to a matrix polynomial generalization of that theorem obtained by Gohberg and Lerer [5]. To prepare for the theorem, we review some facts about zeros of determinants of holomorphic matrix-valued functions.

Let f be a $p \times p$ matrix-valued function that is holomorphic on a connected open set Ω in $\mathbb{C}$ such that $\det f(\lambda)$ is not identically zero on Ω. For any λ_0 in Ω such that $\det f(\lambda_0) = 0$, the *multiplicity* of λ_0 is the order of λ_0 as a zero of the holomorphic function $\lambda \mapsto \det f(\lambda)$. If Γ is a Cauchy contour in Ω such that $\det f$ has no zero on Γ, then $\det f$ has only a finite number of zeros inside Γ, and we denote the sum of their multiplicities by $m(\Gamma, f)$. It can be shown that

$$m(\Gamma, f) = \operatorname{tr}\left(\frac{1}{2\pi i} \int_\Gamma f'(z) f(z)^{-1}\, dz \right). \tag{2.16}$$

See [4, Sect. XI.9] for a proof, along with a proof of the following generalization of Rouché's Theorem.

THEOREM 2.2. *Let f and h be $p \times p$ matrix-valued functions holomorphic on Ω, with $f(z)$ invertible for $z \in \Gamma$. If*

$$\|f(z)^{-1} h(z)\| < 1 \qquad (z \in \Gamma)$$

then $f(z) + h(z)$ is invertible for $z \in \Gamma$, and

$$m(\Gamma, f + h) = m(\Gamma, f).$$

In the next theorem the zeros and eigenvalues are counted according to their multiplicities.

THEOREM 2.3. *Suppose that a, b, and G are as in Theorem 2.1 and that*

$$\begin{pmatrix} I & G \\ G^* & I \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} e_1 \\ 0 \end{pmatrix}.$$

Let $\alpha(z) = \alpha_0 + \alpha_1 z + \cdots + \alpha_n z^n$, $\beta(z) = \beta_0 + \beta_{-1} z^{-1} + \cdots + \beta_{-n} z^{-n}$, and

$$f(z) = \alpha_0 z^{2n+1} + \cdots + \alpha_n z^{n+1} + \beta_{-n} z^n + \cdots + \beta_0.$$

a. *If* α_0 *is positive definite, then* $\alpha(z)$ *and* $f(z)$ *are invertible for all* z *on the unit circle, and the following numbers are equal*:

(i) *the number of zeros of* $\det \alpha$ *inside the unit circle*;

(ii) *the number of zeros of* $\det f$ *outside the unit circle*; *and*

(iii) *the number of negative eigenvalues of*

$$\begin{pmatrix} I & G_1 \\ G_1^* & I \end{pmatrix}, \qquad \text{where } G_1 = \begin{pmatrix} g_n & \cdots & g_1 \\ & \ddots & \vdots \\ 0 & & g_n \end{pmatrix}.$$

b. *If* α_0 *is negative definite, then* $\beta(z)$ *and* $f(z)$ *are invertible for all* z *on the unit circle, and the following number are equal*:

(i) *the number of zeros of* $\det \beta$ *outside the unit circle*;

(ii) *the number of zeros of* $\det f$ *inside the unit circle*; *and*

(iii) *the number of negative eigenvalues of* $\begin{pmatrix} I & G_1 \\ G_1^* & I \end{pmatrix}$.

Proof. Assume that α_0 is positive definite. Then (2.3) implies that

$$\alpha(z)^* \alpha(z) = \alpha_0 + \beta(z)^* \beta(z) \qquad (|z| = 1).$$

Thus $\alpha(z)^*\alpha(z)$ is positive definite and hence invertible for $|z| = 1$. Therefore $\alpha(z)$ is invertible for $|z| = 1$. Suppose that for some z on the unit circle, $f(z)$ is not invertible and let x be a nonzero vector such that $f(z)x = 0$. Since

$$f(z) = z^{2n+1} \alpha\left(\frac{1}{z}\right) + \beta\left(\frac{1}{z}\right)$$

we have

$$\beta\left(\frac{1}{z}\right) x = -z^{2n+1} \alpha\left(\frac{1}{z}\right) x.$$

Using (2.3), we have

$$\begin{aligned} x^*\alpha_0 x &= x^*\alpha\left(\frac{1}{z}\right)^* \alpha\left(\frac{1}{z}\right) x - x^*\beta\left(\frac{1}{z}\right)^* \beta\left(\frac{1}{z}\right) x \\ &= x^*\alpha\left(\frac{1}{z}\right)^* \alpha\left(\frac{1}{z}\right) x - x^* z^{-(2n+1)} \alpha\left(\frac{1}{z}\right)^* z^{2n+1} \alpha\left(\frac{1}{z}\right) x = 0, \end{aligned}$$

which contradicts the fact that α_0 is positive definite. Therefore $f(z)$ is invertible for all z on the unit circle. Write

$$f(z) = W(z) + S(z), \tag{2.17}$$

where

$$W(z) = \alpha_0 z^{2n+1} + \cdots + \alpha_n z^{n+1}$$

and

$$S(z) = \beta_{-n} z^n + \cdots + \beta_0.$$

For $|z| = 1$,

$$W(z) = z^{2n+1} \alpha\left(\frac{1}{z}\right) \tag{2.18}$$

and

$$S(z) = \beta\left(\frac{1}{z}\right)$$

so that (2.3) implies

$$W(z)^* W(z) - S(z)^* S(z) = \alpha_0. \tag{2.19}$$

Thus

$$I - \left(S(z)W(z)^{-1}\right)^* \left(S(z)W(z)^{-1}\right) = \left(W(z)^{-1}\right)^* \alpha_0 W(z)^{-1}.$$

Since the right side is positive definite, it follows that

$$\|S(z)W(z)^{-1}\| < 1 \qquad (|z| = 1). \tag{2.20}$$

By Theorem 2.2, (2.17) and (2.20) imply that

$$m(\mathbb{T}, f) = m(\mathbb{T}, W). \tag{2.21}$$

Next, we claim that

$$m(\mathbb{T}, W) = (2n+1)p - m(\mathbb{T}, \alpha). \tag{2.22}$$

We use (2.16) with Γ the unit circle. From (2.18) for $|z| = 1$, we find that

$$W'(z)W(z)^{-1} = \frac{2n+1}{z} I_p - \frac{1}{z^2}\alpha'\left(\frac{1}{z}\right)\alpha\left(\frac{1}{z}\right)^{-1}.$$

Using (2.16), we have

$$\begin{aligned} m(\mathbb{T}, W) &= \operatorname{tr}\left(\frac{1}{2\pi i}\int_{\mathbb{T}} \frac{2n+1}{z} I_p\, dz\right) - \operatorname{tr}\left(\frac{1}{2\pi i}\int_{\mathbb{T}} \frac{1}{z^2}\alpha'\left(\frac{1}{z}\right)\alpha\left(\frac{1}{z}\right)^{-1} dz\right) \\ &= (2n+1)p - \operatorname{tr}\left(\frac{1}{2\pi i}\int_0^{2\pi} e^{-2it}\alpha'(e^{-it})\alpha(e^{-it})^{-1} ie^{it}\, dt\right) \\ &= (2n+1)p - \operatorname{tr}\left(\frac{1}{2\pi i}\int_0^{2\pi} \alpha'(e^{iu})\alpha(e^{iu})^{-1} ie^{iu}\, du\right) \\ &= (2n+1)p - m(\mathbb{T}, \alpha). \end{aligned}$$

This proves (2.22), and (2.21) shows that

$$m(\mathbb{T}, \alpha) = (2n+1)p - m(\mathbb{T}, f). \tag{2.23}$$

Let $h(z) = f(z) - \alpha_0 z^{2n+1}$, and let Γ be a circle about $z = 0$ that encloses all the zeros of $\det f$ and is of sufficiently large radius that

$$\left\|\left(\alpha_0 z^{2n+1}\right)^{-1} h(z)\right\| < 1 \qquad (z \in \Gamma).$$

By Theorem 2.2, the function $z \mapsto \det \alpha_0 z^{2n+1}$ and $\det f$ have the same number of zeros inside Γ. Since α_0 is invertible and

$$\det\left(\alpha_0 z^{2n+1}\right) = (\det \alpha_0)\det\left(z^{2n+1} I_p\right) = (\det \alpha_0) z^{(2n+1)p},$$

both $\det \alpha_0 z^{2n+1}$ and $\det f$ have $(2n+1)p$ zeros, counting multiplicities. Consequently, the right side of the equality in (2.23) is the number of zeros of $\det f$ outside the unit circle, which shows that the number in (i) and (ii) of part (a) of the theorem are equal. By Theorem 2.1 in [5], the number in (ii) equals the number of negative eigenvalues of the matrix obtained by deleting the last (block) column and row of

$$\begin{pmatrix} I & G \\ G^* & I \end{pmatrix}.$$

By repartitioning the resulting matrix in the block form

$$\begin{array}{c} \\ n \\ 1 \\ n \end{array}\begin{array}{c} \begin{array}{ccc} n & 1 & n \end{array} \\ \begin{pmatrix} I & 0 & G_1 \\ 0 & I & 0 \\ G_1^* & 0 & I \end{pmatrix} \end{array}$$

it is clear that this resulting matrix and the matrix in (iii) of part (a) have the same number of negative eigenvalues. Thus the numbers in (ii) and (iii) part (a) are equal.

Now suppose that α_0 is negative definite. By (2.3),

$$\beta^*(z)\beta(z) = \alpha(z)^*\alpha(z) - \alpha_0 > 0$$

so that $\beta(z)$ is invertible for $|z| = 1$. As before, $f(z)$ is invertible for $|z| = 1$ also. By (2.19)

$$S(z)^* S(z) - W(z)^* W(z) = -\alpha_0$$

so that

$$I - \left(W(z)S(z)^{-1}\right)^*\left(W(z)S(z)^{-1}\right) = \left(S(z)^{-1}\right)^*(-\alpha_0)S(z)^{-1} > 0.$$

Therefore

$$\|W(z)S(z)^{-1}\| < 1. \tag{2.24}$$

Hence, by Theorem 2.2, it follows from (2.17) and (2.24) that

$$m(\mathbb{T}, f) = m(\mathbb{T}, S).$$

Since $S(z) = \beta(1/z)$, we conclude that $m(\mathbb{T}, f)$ equals the number of zeros of det β outside the unit circle. Thus (i) and (ii) in part (b) of the theorem are equal. By [5, Theorem 2.1], (i), (ii), and (iii) are equal. ∎

In the next theorem we give necessary and sufficient conditions for the existence of a block upper-triangular matrix G satisfying not only

$$\begin{pmatrix} I & G \\ G^* & I \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} e_1 \\ 0 \end{pmatrix} \tag{2.25}$$

for given a and b, but also

$$\begin{pmatrix} I & G \\ G^* & I \end{pmatrix}\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ \hat{e}_1 \end{pmatrix} \tag{2.26}$$

for given c and d, where $\hat{e}_1 = (0, \dots, 0, I)^T$. We let $c = (\gamma_0, \dots, \gamma_n)^T$ and $d = (\delta_{-n}, \dots, \delta_0)^T$ and define

$$\gamma(z) = \gamma_0 + \gamma_1 z + \cdots + \gamma_n z^n$$

and

$$\delta(z) = \delta_0 + \delta_{-1} z^{-1} + \cdots + \delta_{-n} z^{-n}.$$

THEOREM 2.4. *Let* $a = (\alpha_0, \dots, \alpha_n)^T$, $b = (\beta_{-n}, \dots, \beta_0)^T$, $c = (\gamma_0, \dots, \gamma_n)^T$, $d = (\delta_{-n}, \dots, \delta_0)^T$ *be given block vectors with* α_0 *and* δ_0 *invertible. Then there is a block upper-triangular matrix G such that* (2.25) *and* (2.26) *are satisfied if and only if* α_0 *and* δ_0 *are self-adjoint and the following three conditions hold*:

(i)

$$\begin{pmatrix} \alpha_0^* & \cdots & \alpha_n^* \\ & \ddots & \vdots \\ 0 & & \alpha_0^* \end{pmatrix}\begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix} - \begin{pmatrix} \beta_{-n}^* & \cdots & \beta_0^* \\ & \ddots & \vdots \\ 0 & & \beta_{-n}^* \end{pmatrix}\begin{pmatrix} \beta_{-n} \\ \vdots \\ \beta_0 \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

(ii)

$$\begin{pmatrix} \delta_0^* & & 0 \\ \vdots & \ddots & \\ \delta_{-n}^* & \cdots & \delta_0^* \end{pmatrix}\begin{pmatrix} \delta_{-n} \\ \vdots \\ \delta_0 \end{pmatrix} - \begin{pmatrix} \gamma_n^* & & 0 \\ \vdots & \ddots & \\ \gamma_0^* & \cdots & \gamma_n^* \end{pmatrix}\begin{pmatrix} \gamma_0 \\ \vdots \\ \gamma_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \delta_0 \end{pmatrix}$$

(iii)

$$\begin{pmatrix} \alpha_0^* & \cdots & \alpha_n^* \\ & \ddots & \vdots \\ 0 & & \alpha_0^* \end{pmatrix}\begin{pmatrix} \gamma_n & \cdots & \gamma_0 \\ & \ddots & \vdots \\ 0 & & \gamma_n \end{pmatrix}$$
$$= \begin{pmatrix} \beta_{-n}^* & \cdots & \beta_0^* \\ & \ddots & \vdots \\ 0 & & \beta_{-n}^* \end{pmatrix}\begin{pmatrix} \delta_0 & \cdots & \delta_{-n} \\ & \ddots & \vdots \\ 0 & & \delta_0 \end{pmatrix}.$$

The matrix G is unique and is given by

$$G = -\begin{pmatrix} \alpha_0^* & \cdots & \alpha_n^* \\ & \ddots & \vdots \\ 0 & & \alpha_0^* \end{pmatrix}^{-1} \begin{pmatrix} \beta_{-n}^* & \cdots & \beta_0^* \\ & \ddots & \vdots \\ 0 & & \beta_{-n}^* \end{pmatrix}$$

$$= -\begin{pmatrix} \gamma_n & \cdots & \gamma_0 \\ & \ddots & \vdots \\ 0 & & \gamma_n \end{pmatrix} \begin{pmatrix} \delta_0 & \cdots & \delta_{-n} \\ & \ddots & \vdots \\ 0 & & \delta_0 \end{pmatrix}^{-1}.$$

The equations in (i)–(iii) *are equiavlent to*

(i′) $\alpha(z)^*\alpha(z) - \beta(z)^*\beta(z) = \alpha_0$ $(|z| = 1)$
(ii′) $\delta(z)^*\delta(z) - \gamma(z)^*\gamma(z) = \delta_0$ $(|z| = 1)$
(iii′) *The coefficients of the nonnegative powers of z are the same for $\alpha(z)^*\gamma(z)$ and $\beta(z)^*\delta(z)$.*

Proof. Suppose there exists such a G. By Theorem 2.1 we know that α_0 is self-adjoint, (i) holds, and (i) can be rewritten as (i′). We will obtain (ii) and (ii′) in an analogous way. Equation (2.26) implies

$$c + Gd = 0$$

$$c^*G + d^* = (\hat{e}_1)^*.$$

Writing out the entries, we have

$$\begin{pmatrix} \gamma_0 \\ \vdots \\ \gamma_n \end{pmatrix} + \begin{pmatrix} g_n & \cdots & g_0 \\ & \ddots & \vdots \\ 0 & & g_n \end{pmatrix} \begin{pmatrix} \delta_{-n} \\ \vdots \\ \delta_0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} \gamma_0^* & \cdots & \gamma_n^* \end{pmatrix} \begin{pmatrix} g_n & \cdots & g_0 \\ & \ddots & \vdots \\ 0 & & g_n \end{pmatrix} + \begin{pmatrix} \delta_{-n}^* & \cdots & \delta_0^* \end{pmatrix}$$

$$= \begin{pmatrix} 0 & \cdots & 0 & I \end{pmatrix}.$$

These imply

$$\begin{pmatrix} g_n & \cdots & g_0 \\ & \ddots & \vdots \\ 0 & & g_n \end{pmatrix}\begin{pmatrix} \delta_{-n} & \cdots & \delta_0 \\ \vdots & \cdot & \\ \delta_0 & & 0 \end{pmatrix} + \begin{pmatrix} \gamma_0 & \cdots & \gamma_n \\ \vdots & \cdot & \\ \gamma_n & & 0 \end{pmatrix} = 0 \quad (2.27)$$

and

$$\begin{pmatrix} \gamma_0^* & \cdots & \gamma_n^* \\ & \ddots & \vdots \\ 0 & & \gamma_0^* \end{pmatrix}\begin{pmatrix} g_n & \cdots & g_0 \\ & \ddots & \vdots \\ 0 & & g_n \end{pmatrix} = \begin{pmatrix} -\delta_{-n}^* & \cdots & -\delta_{-1}^* & I-\delta_0^* \\ & & & -\delta_{-1}^* \\ & & \ddots & \vdots \\ 0 & & & -\delta_{-n}^* \end{pmatrix}.$$

These imply in turn that

$$-\begin{pmatrix} \gamma_0^* & \cdots & \gamma_n^* \\ & \ddots & \vdots \\ 0 & & \gamma_0^* \end{pmatrix}\begin{pmatrix} \gamma_0 & \cdots & \gamma_n \\ \vdots & \cdot & \\ \gamma_n & & 0 \end{pmatrix}$$

$$-\begin{pmatrix} -\delta_{-n}^* & \cdots & -\delta_{-1}^* & I-\delta_0^* \\ & & & -\delta_{-1}^* \\ & & \ddots & \vdots \\ 0 & & & -\delta_{-n}^* \end{pmatrix}\begin{pmatrix} \delta_{-n} & \cdots & \delta_0 \\ \vdots & \cdot & \\ \delta_0 & & 0 \end{pmatrix} = 0$$

so that

$$\begin{pmatrix} \delta_{-n}^* & \cdots & \delta_0^* \\ & \ddots & \vdots \\ 0 & & \delta_{-n}^* \end{pmatrix}\begin{pmatrix} \delta_{-n} & \cdots & \delta_0 \\ \vdots & \cdot & \\ \delta_0 & & 0 \end{pmatrix} - \begin{pmatrix} \gamma_0^* & \cdots & \gamma_n^* \\ & \ddots & \vdots \\ 0 & & \gamma_0^* \end{pmatrix}\begin{pmatrix} \gamma_0 & \cdots & \gamma_n \\ \vdots & \cdot & \\ \gamma_n & & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \delta_0 & \cdots & 0 \\ 0 & & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

This equation implies that δ_0 is self-adjoint. Taking the adjoint of this equation and equating the first columns, we obtain

$$\begin{pmatrix} \delta_{-n}^* & \cdots & \delta_0^* \\ \vdots & \cdot & \\ \delta_0^* & & 0 \end{pmatrix}\begin{pmatrix} \delta_{-n} \\ \vdots \\ \delta_0 \end{pmatrix} - \begin{pmatrix} \gamma_0^* & \cdots & \gamma_n^* \\ \vdots & \cdot & \\ \gamma_n^* & & 0 \end{pmatrix}\begin{pmatrix} \gamma_0 \\ \vdots \\ \gamma_n \end{pmatrix} = \begin{pmatrix} \delta_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Equation (ii) follows by reversing the rows in the equation above. It is easy to see that (ii) implies (ii′). From Theorem 2.1, we have

$$G = -\begin{pmatrix} \alpha_0^* & \cdots & \alpha_n^* \\ & \ddots & \vdots \\ 0 & & \alpha_0^* \end{pmatrix}^{-1}\begin{pmatrix} \beta_{-n}^* & \cdots & \beta_0^* \\ & \ddots & \vdots \\ 0 & & \beta_{-n}^* \end{pmatrix} \tag{2.28}$$

and from (2.27) it follows that

$$\begin{aligned} G &= -\begin{pmatrix} \gamma_0 & \cdots & \gamma_n \\ \vdots & \cdot & \\ \gamma_n & & 0 \end{pmatrix}\begin{pmatrix} \delta_{-n} & \cdots & \delta_0 \\ \vdots & \cdot & \\ \delta_0 & & 0 \end{pmatrix}^{-1} \\ &= -\begin{pmatrix} \gamma_n & \cdots & \gamma_0 \\ & \ddots & \vdots \\ 0 & & \gamma_n \end{pmatrix}\begin{pmatrix} \delta_0 & \cdots & \delta_{-n} \\ & \ddots & \vdots \\ 0 & & \delta_0 \end{pmatrix}^{-1}. \end{aligned} \tag{2.29}$$

Equations (2.28) and (2.29) imply (iii). A simple calculation shows that the entries in the first row of the product on the left (respectively, right) side of (iii) are the coefficients of the nonnegative powers of z in the product $\alpha(z)^*\gamma(z)$ (respectively, $\beta(z)^*\delta(z)$). Thus (iii) is equivalent to (iii′).

Now suppose that conditions (i)–(iii) are satisfied, and define G by (2.28). Then G is a block upper-triangular Toeplitz matrix, and by Theorem 2.1,

$$\begin{pmatrix} I & G \\ G^* & I \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} e_1 \\ 0 \end{pmatrix}.$$

By (iii), G is also given by (2.29), which implies that

$$G\begin{pmatrix} \delta_0 & \cdots & \delta_{-n} \\ & \ddots & \vdots \\ 0 & & \delta_0 \end{pmatrix} + \begin{pmatrix} \gamma_n & \cdots & \gamma_0 \\ & \ddots & \vdots \\ 0 & & \gamma_n \end{pmatrix} = 0.$$

Taking the last columns, we have

$$Gd + c = 0. \tag{2.30}$$

Using (2.29) and (ii), we find that

$$G^*c + d = -\begin{pmatrix} \delta_0^* & & 0 \\ \vdots & \ddots & \\ \delta_{-n}^* & \cdots & \delta_0^* \end{pmatrix}^{-1} \begin{pmatrix} \gamma_n^* & & 0 \\ \vdots & \ddots & \\ \gamma_0^* & \cdots & \gamma_n^* \end{pmatrix} \begin{pmatrix} \gamma_0 \\ \vdots \\ \gamma_n \end{pmatrix} + \begin{pmatrix} \delta_{-n} \\ \vdots \\ \delta_0 \end{pmatrix}$$

$$= -\begin{pmatrix} \delta_0^* & & 0 \\ \vdots & \ddots & \\ \delta_{-n}^* & \cdots & \delta_0^* \end{pmatrix}^{-1} \left[\begin{pmatrix} \delta_0^* & & 0 \\ \vdots & \ddots & \\ \delta_{-n}^* & \cdots & \delta_0^* \end{pmatrix} \begin{pmatrix} \delta_{-n} \\ \vdots \\ \delta_0 \end{pmatrix} - \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \delta_0 \end{pmatrix} \right]$$

$$+ \begin{pmatrix} \delta_{-n} \\ \vdots \\ \delta_0 \end{pmatrix}$$

$$= \begin{pmatrix} \delta_0^* & & 0 \\ \vdots & \ddots & \\ \delta_{-n}^* & \cdots & \delta_0^* \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \delta_0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ I \end{pmatrix} = \hat{e}_1.$$

Combining this result with (2.30) yields

$$\begin{pmatrix} I & G \\ G^* & I \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ \hat{e}_1 \end{pmatrix}.$$

Thus G is the desired matrix. By (2.28) and (2.29), G is unique. ∎

3. AN EXAMPLE

Let g be any $p \times p$ complex matrix such that $I - gg^*$ is invertible, and let G be the $(n+1) \times (n+1)$ block diagonal matrix

$$G = \begin{pmatrix} g & 0 & \cdots & 0 \\ 0 & g & & 0 \\ \vdots & & \ddots & \\ 0 & & \cdots & g \end{pmatrix}.$$

Suppose a and b satisfy

$$\begin{pmatrix} I & G \\ G^* & I \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} e_1 \\ 0 \end{pmatrix}.$$

Then

$$b = -G^*a, \tag{3.1}$$

and so from the equation $a + Gb = e_1$, we obtain $(I - GG^*)a = e_1$, and hence

$$a = (I - GG^*)^{-1}e_1, \tag{3.2}$$

where $I - GG^*$ is a block diagonal matrix with $I - gg^*$ in each diagonal position. Let $a = (\alpha_0, \dots, \alpha_n)^T$ and $b = (\beta_{-n}, \dots, \beta_0)^T$. Then (3.2) shows that $\alpha_0 = (I - gg^*)^{-1}$ and $\alpha_1 = \cdots = \alpha_n = 0$. From (3.1), $\beta_{-n} = -g^*(I - gg^*)^{-1}$ and $\beta_{-n+1} = \cdots = \beta_0 = 0$. Thus the functions α and β have the form

$$\alpha(z) = \alpha_0 = (I - gg^*)^{-1}$$

and

$$\beta(z) = \beta_{-n}z^{-n} = -g^*(I - gg^*)^{-1}z^{-n}.$$

Suppose that α_0 is positive definite. This happens precisely when $\|g\| < 1$. Let G_1 be the $n \times n$ block diagonal matrix with g in each diagonal position. Then the matrix factorization

$$T_1 = \begin{pmatrix} I & G_1 \\ G_1^* & I \end{pmatrix} = \begin{pmatrix} I & G_1 \\ 0 & I \end{pmatrix}\begin{pmatrix} I - G_1G_1^* & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} I & 0 \\ G_1^* & I \end{pmatrix} \tag{3.3}$$

shows that T_1 is positive definite. Thus the number of negative eigenvalues of T_1 equals the number of zeros of $\det \alpha$ inside the unit circle (because both numbers are zero), as described in Theorem 2.3a.

Next, suppose that α_0 is negative definite. (This is the case precisely when g is invertible and $\|g^{-1}\| < 1$.) Then $I - G_1G_1^*$ has np negative eigenvalues, and so does T_1, by (3.3). This number matches the number of zeros of $\det \beta$ outside the unit circle, as described in Theorem 2.3b.

Finally, suppose that α_0 is indefinite but invertible. The calculations above show that α is still a constant function, with det α having no zeros inside the unit circle, and det β has at least pn zeros outside the unit disk. However, the number of negative eigenvalues of T_1 lies strictly between 0 and pn and hence cannot match the number of zeros of either det α or det β.

REFERENCES

1 P. Duren, *Theory of H^p Spaces*, Academic Press, New York, 1970.
2 R. L. Ellis and I. Gohberg, Orthogonal systems related to infinite Hankel matrices, *J. Funct. Anal.* 109:155–198 (1992).
3 R. L. Ellis, I. Gohberg, and D. C. Lay, Infinite analogues of block Toeplitz matrices and related orthogonal functions, *Integral Equations and Operator Theory* 22: 375–419 (1995).
4 I. Gohberg, S. Goldberg, and M. A. Kaashoek, *Classes of Linear Operators*, Vol. I, *Operator Theory: Adv. Appl.*, Vol. 49, Birkhäuser, Basel, 1990.
5 I. Gohberg and L. Lerer, Matrix generalizations of M. G. Krein theorems on orthogonal polynomials, *Operator Theory: Adv. Appl.*, Vol. 34, pp. 137–202, Birkhäuser, Basel, 1988.
6 M. G. Krein, On the distribution of the roots of polynomials which are orthogonal on the unit circle with respect to an alternating weight, *Teor. Funktsiĭ Funkstional Anal. i Prilozhen* 2:131–137 (1966). [In Russian]

Received 6 October 1994; final manuscript accepted 20 March 1995